

SURJECTIVE KUIPER ISOMETRIES

GYÖRGY PÁL GEHÉR

ABSTRACT. Characterisations of surjective isometries with respect to the Kuiper distance on three classes of Borel probability measures of \mathbb{R} (or equivalently, probability distribution functions) are presented here. These classes are the set of continuous, absolute continuous and general measures.

1. INTRODUCTION AND STATEMENTS OF THE RESULTS

The famous Banach–Stone theorem characterises surjective linear isometries between Banach spaces of continuous complex-valued functions on compact Hausdorff spaces equipped with the supremum norm. Motivated by this theorem, G. Dolinar and L. Molnár described surjective isometries of probability distribution functions with respect to the Kolmogorov–Smirnov distance in their joint paper [4]. It is important to note that the space of probability distribution functions is not a linear space, therefore there is no point in considering linearity of these maps. Though, this space is a convex subset of a linear space, i.e. the space of all real-valued measures on \mathbb{R} , therefore invariance of convex combinations could be considered. But on the contrary, in [4] the authors considered general surjective isometries, and the invariance of convex combinations under these transformations was not an assumption but a conclusion.

Later, L. Molnár continued this project in three further publications [7, 8, 9]. The present paper is a contribution to this line of research. Namely, we will consider the so-called Kuiper distance and describe the structure of general surjective isometries on three classes of probability distribution functions. However, unlike in the before mentioned papers, we will prefer the language of Borel probability measures instead of distribution functions. We point out that as in the papers [4, 7, 8, 9], the invariance of convex combinations under these transformations will be a conclusion.

The space of all Borel probability measures will be denoted by $P(\mathbb{R})$. We call a $\mu \in P(\mathbb{R})$ continuous if we have $\mu(\{x\}) = 0$ for every $x \in \mathbb{R}$, and absolute continuous if it is absolutely continuous with respect to the Gaussian measure (or equivalently with respect to the usual length measure m on \mathbb{R} , though, we have to keep in mind that $m \notin P(\mathbb{R})$). The class of all continuous and absolutely continuous Borel probability measures on \mathbb{R} will be denoted by $P_c(\mathbb{R})$ and $P_{ac}(\mathbb{R})$, respectively. The symbol $\mathcal{B}_{\mathbb{R}}$ will stand for the set of all Borel subsets of \mathbb{R} .

The distribution function of $\mu \in P(\mathbb{R})$ is usually defined by $f_{\mu}(t) := \mu((-\infty, t])$ ($t \in \mathbb{R}$). It is well-known that f_{μ} is monotone increasing, continuous from the right, and satisfies $\lim_{t \rightarrow \infty} f_{\mu}(t) = 1$ and $\lim_{t \rightarrow -\infty} f_{\mu}(t) = 0$. The Kolmogorov–Smirnov

2010 *Mathematics Subject Classification*. Primary: 47B49, 54E40, Secondary: 47B38, 60B10.

Key words and phrases. Borel probability measures on \mathbb{R} , Kuiper distance, surjective isometries.

distance on $P(\mathbb{R})$ is given by

$$d_{KS}(\mu, \nu) = \sup_{t \in \mathbb{R}} |f_\mu(t) - f_\nu(t)| = \sup_{t \in \mathbb{R}} |\mu((-\infty, t]) - \nu((-\infty, t])|,$$

and the so-called total variation distance or statistical distance is

$$d_{TV}(\mu, \nu) = \sup \{ |\mu(B) - \nu(B)| : B \in \mathcal{B}_{\mathbb{R}} \}.$$

Let \mathcal{I} denote the set of all non-degenerate intervals of \mathbb{R} , i.e.

$$\mathcal{I} = \{I \subseteq \mathbb{R} : \#I > 1 \text{ and } I \text{ is connected}\}.$$

The set of all (possibly degenerate) intervals of \mathbb{R} will be denoted by \mathcal{I}_0 , i.e.

$$\mathcal{I}_0 = \mathcal{I} \cup \{\{x\} : x \in \mathbb{R}\} = \{I \subseteq \mathbb{R} : I \text{ is connected and } I \neq \emptyset\}.$$

The Kuiper metric is given by the following formula, where the second and third equations are easy to see:

$$\begin{aligned} d_{Ku}(\mu, \nu) &:= \sup_{t \in \mathbb{R}} (f_\mu(t) - f_\nu(t)) + \sup_{t \in \mathbb{R}} (f_\nu(t) - f_\mu(t)) \\ &= \sup \{ |\mu(I) - \nu(I)| : I \in \mathcal{I} \} = \sup \{ |\mu(I) - \nu(I)| : I \in \mathcal{I}, I \text{ is bounded} \}. \end{aligned} \quad (1.1)$$

This metric is a natural modified version of the above mentioned two distances. Obviously, we have $d_{Ku}(\mu, \nu) \leq 1$ ($\mu, \nu \in P(\mathbb{R})$). We will show in the next section that instead of supremum we can take maximum, if I runs through \mathcal{I}_0 instead of \mathcal{I} . This metric was defined by N.H. Kuiper in [6], and it seems that it is more useful in statistics than the Kolmogorov–Smirnov distance ([1, pp. 39]). It is also important because of the Kuiper density problem ([2]).

The goal of this paper is to present a characterisation of surjective isometries of $P_c(\mathbb{R})$, $P_{ac}(\mathbb{R})$ and $P(\mathbb{R})$ with respect to the Kuiper distance. We remark that this question was posed by Molnár in a personal conversation. Let $A \in \mathcal{B}_{\mathbb{R}}$ and $g : A \rightarrow \mathbb{R}$ be an injective function which transforms Borel sets into Borel sets, i.e. $g(B) = g(B \cap A) \in \mathcal{B}_{\mathbb{R}}$ ($B \in \mathcal{B}_{\mathbb{R}}$). If $\mu \in P(\mathbb{R})$, then by $\mu \circ g$ we mean the (not necessarily probability) Borel measure defined by

$$(\mu \circ g)(B) = \mu(g(B)) \quad (B \in \mathcal{B}_{\mathbb{R}}).$$

Clearly, in the special case when $g : A \rightarrow \mathbb{R}$ is surjective, we have $\mu \circ g \in P(\mathbb{R})$. We also point out that if $\mu \in P_c(\mathbb{R})$ and $\#(\mathbb{R} \setminus g(A)) = 1$, then $\mu \circ g \in P_c(\mathbb{R})$.

For every $x \in \mathbb{R}$ let us define the function

$$r_x : \mathbb{R} \setminus \{x\} \rightarrow \mathbb{R}, \quad r_x(t) = \frac{1}{t - x};$$

and let $r_\infty : \mathbb{R} \rightarrow \mathbb{R}$ be the identity function.

Now, we formulate our main results on surjective Kuiper isometries.

Theorem 1.1. *Let $\phi : P_c(\mathbb{R}) \rightarrow P_c(\mathbb{R})$ be a surjective transformation which is an isometry with respect to d_{Ku} , i.e. we have*

$$d_{Ku}(\mu, \nu) = d_{Ku}(\phi(\mu), \phi(\nu)) \quad (\mu, \nu \in P_c(\mathbb{R})).$$

Then there exists a homeomorphism $g : \mathbb{R} \rightarrow \mathbb{R}$ and an $x \in \mathbb{R} \cup \{\infty\}$ such that ϕ has the following form:

$$\phi(\mu) = \mu \circ (g \circ r_x) \quad (\mu \in P_c(\mathbb{R})). \quad (1.2)$$

Moreover, transformations of the above form are all surjective isometries with respect to the Kuiper metric.

Theorem 1.2. *Let $\phi: P_{ac}(\mathbb{R}) \rightarrow P_{ac}(\mathbb{R})$ be a surjective map which is an isometry with respect to the Kuiper metric, i.e. we have*

$$d_{Ku}(\mu, \nu) = d_{Ku}(\phi(\mu), \phi(\nu)) \quad (\mu, \nu \in P_{ac}(\mathbb{R})).$$

Then there exists a bijective function $g: \mathbb{R} \rightarrow \mathbb{R}$ such that g and g^{-1} are locally (i.e. on every compact interval) absolutely continuous, and an $x \in \mathbb{R} \cup \{\infty\}$ such that we have

$$\phi(\mu) = \mu \circ (g \circ r_x) \quad (\mu \in P_{ac}(\mathbb{R})).$$

Moreover, every transformation with this form is a surjective Kuiper isometry on $P_{ac}(\mathbb{R})$.

As a consequence of the above two theorems we have that every surjective Kuiper isometry on $P_{ac}(\mathbb{R})$ can be extended to a surjective Kuiper isometry on $P_c(\mathbb{R})$. But on the contrary, as we shall see from the next theorem, a similar conclusion does not hold for the classes $P(\mathbb{R})$ and $P_c(\mathbb{R})$. The reason is the following: unlike on $P_c(\mathbb{R})$, all surjective Kuiper isometries on $P(\mathbb{R})$ transform measures with compact support to measures of the same type.

Theorem 1.3. *Let $\phi: P(\mathbb{R}) \rightarrow P(\mathbb{R})$ be a surjective Kuiper isometry, i.e. we have*

$$d_{Ku}(\mu, \nu) = d_{Ku}(\phi(\mu), \phi(\nu)) \quad (\mu, \nu \in P(\mathbb{R})).$$

Then there exists a homeomorphism $g: \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\phi(\mu) = \mu \circ g \quad (\mu \in P(\mathbb{R})).$$

Moreover, every transformation of this form is a surjective Kuiper isometry on $P(\mathbb{R})$.

The first two theorems will be proven in Section 2. Our method will be to transform our problem to another one which considers Borel probability measures on the unit circle \mathbb{T} of \mathbb{R}^2 . This will help us to avoid some technical difficulties. The last result will be verified in Section 3. In order to give that characterisation, our first step will be to establish a metric characterisation of Dirac measures in a similar way as it was done in [4] and [8]. After that we shall utilise Theorem 1.1.

2. PROOFS OF THE CONTINUOUS CASES

We begin with proving that in (1.1) we can take maximum if I runs through \mathcal{I}_0 .

Lemma 2.1. *For every $\mu, \nu \in P(\mathbb{R})$ we have the following equation:*

$$d_{Ku}(\mu, \nu) = \max\{|\mu(I) - \nu(I)|: I \in \mathcal{I}_0\}. \quad (2.1)$$

Proof. If $d_{Ku}(\mu, \nu) = 0$, then we may choose $I = \{x_0\}$ with $\mu(\{x_0\}) = \nu(\{x_0\}) = 0$. Therefore we may assume throughout the proof that $d_{Ku}(\mu, \nu) > 0$.

First, let us assume that $f_\mu \leq f_\nu$ (the $f_\mu \geq f_\nu$ case is similar). Then we clearly have

$$d_{Ku}(\mu, \nu) = \sup_{t \in \mathbb{R}} (f_\nu(t) - f_\mu(t)).$$

Since the limits of the distribution functions at $-\infty$ and ∞ are finite and they coincide, we can find a bounded sequence $\{t_n\}_{n=1}^\infty$ such that $\lim_{n \rightarrow \infty} f_\nu(t_n) - f_\mu(t_n) = d_{Ku}(\mu, \nu)$. Since every bounded sequence has a convergent subsequence, we may

assume that our sequence was convergent. Let $t_0 := \lim_{n \rightarrow \infty} t_n$. We can also suppose that we have either $t_n \geq t_0$ ($n \in \mathbb{N}$) or $t_n < t_0$ ($n \in \mathbb{N}$). If the first possibility occurs, then we obtain

$$d_{Ku}(\mu, \nu) = \lim_{n \rightarrow \infty} f_\nu(t_n) - f_\mu(t_n) = f_\nu(t_0) - f_\mu(t_0) = \nu((-\infty, t_0]) - \mu((-\infty, t_0]).$$

For the second one, we infer

$$d_{Ku}(\mu, \nu) = \lim_{n \rightarrow \infty} f_\nu(t_n) - f_\mu(t_n) = f_\nu(t_0-) - f_\mu(t_0-) = \nu((-\infty, t_0)) - \mu((-\infty, t_0)).$$

Second, we assume that neither $f_\mu \leq f_\nu$ nor $f_\mu \geq f_\nu$ holds. In this case there exist bounded sequences $\{t_n\}_{n=1}^\infty$ and $\{s_n\}_{n=1}^\infty$ such that

$$d_{Ku}(\mu, \nu) = \lim_{n \rightarrow \infty} (f_\mu(t_n) - f_\nu(t_n)) + \lim_{n \rightarrow \infty} (f_\nu(s_n) - f_\mu(s_n)).$$

Like in the previous case, we may assume that these sequences are convergent. Let $s_0 := \lim_{n \rightarrow \infty} s_n$ and $t_0 := \lim_{n \rightarrow \infty} t_n$. We can also suppose that we have either $t_n \geq t_0$ ($n \in \mathbb{N}$) or $t_n < t_0$ ($n \in \mathbb{N}$); and a similar assumption can be made on $\{s_n\}_{n=1}^\infty$. Let us assume that we have $s_n < s_0$ and $t_0 \leq t_n$ ($n \in \mathbb{N}$). Then we deduce the following:

$$d_{Ku}(\mu, \nu) = f_\mu(t_0) - f_\nu(t_0) + f_\nu(s_0-) - f_\mu(s_0-) = \begin{cases} \mu([s_0, t_0]) - \nu([s_0, t_0]) & \text{if } s_0 \leq t_0 \\ \nu((t_0, s_0)) - \mu((t_0, s_0)) & \text{if } t_0 < s_0 \end{cases}.$$

The other cases can be handled similarly. \square

Let $\xi_t = (\cos t, \sin t) \in \mathbb{R}^2$ ($t \in \mathbb{R}$), and let us consider the unit circle $\mathbb{T} = \{\xi_t : t \in [-\pi, \pi)\}$, and the following continuous map:

$$\tau : \mathbb{R} \rightarrow \mathbb{T}, \quad \tau(t) = \xi_{2 \arctan t}.$$

The set of all (possibly degenerate) arcs of \mathbb{T} is denoted by \mathcal{A}_0 . The sets of all non-degenerate arcs, closed (possibly degenerate) arcs, and open arcs will be denoted by the symbols \mathcal{A} , \mathcal{A}^{cl} and \mathcal{A}^{op} , respectively. (Note that none of \mathcal{A} , \mathcal{A}_0 , \mathcal{A}^{cl} or \mathcal{A}^{op} contains the empty set). The spaces of all continuous and absolutely continuous (with respect to the normalised arc-length measure λ on \mathbb{T}) Borel probability measures will be denoted by $P_c(\mathbb{T})$ and $P_{ac}(\mathbb{T})$, respectively. We define the following function:

$$d : P_c(\mathbb{T}) \times P_c(\mathbb{T}) \rightarrow \mathbb{R}, \quad d(\mu, \nu) = \max\{|\mu(A) - \nu(A)| : A \in \mathcal{A}_0\}. \quad (2.2)$$

Clearly, the map τ is a homeomorphism between \mathbb{R} and $\mathbb{T} \setminus \{\xi_\pi\}$. We observe the following:

$$d(\mu, \nu) = d_{Ku}(\mu \circ \tau, \nu \circ \tau) \quad (\mu, \nu \in P_c(\mathbb{T})),$$

where $\mu \circ \tau \in P_c(\mathbb{R})$ is defined by $(\mu \circ \tau)(B) = \mu(\tau(B))$ ($B \in \mathcal{B}_{\mathbb{R}}$). In order to verify this, we observe two properties. First, that we have $d_{Ku}(\mu, \nu) = \max\{|\mu(\mathbb{R} \setminus I) - \nu(\mathbb{R} \setminus I)| : I \in \mathcal{I}_0\}$. Second, that for every $I \in \mathcal{I}_0$ the sets $\tau(I)$ and $\{-1\} \cup \tau(\mathbb{R} \setminus I)$ are arcs of \mathbb{T} , and that for every arc $A \subset \mathbb{T}$ the set $\tau^{-1}(A)$ (inverse image) is either an interval or the complement of an interval. Therefore we conclude that d is a metric on $P_c(\mathbb{T})$. We note that μ is absolutely continuous exactly when $\mu \circ \tau$ is. In fact, this is a consequence of the fact that the derivatives of the tangent and arctangent functions are bounded from below and above by some positive numbers on every compact interval of $(-\frac{\pi}{2}, \frac{\pi}{2})$ and \mathbb{R} , respectively.

In order to prove Theorems 1.1 and 1.2 first, we investigate surjective isometries on $P_c(\mathbb{T})$ and $P_{ac}(\mathbb{T})$ with respect to the metric d . For any $\mu \in P_c(\mathbb{T})$ we will use the following notation:

$$\{\mu\}^c := \{\nu \in P_c(\mathbb{T}) : d(\mu, \nu) = 1\}.$$

If $\mu \in P_{ac}(\mathbb{T})$, then we define

$$\{\mu\}^{ac} := \{\nu \in P_{ac}(\mathbb{T}) : d(\mu, \nu) = 1\}.$$

The closed support of $\mu \in P_c(\mathbb{T})$ (or $\mu \in P(\mathbb{R})$, respectively) is the smallest closed set S_μ of \mathbb{T} (or \mathbb{R} , resp.) such that we have $\mu(S_\mu) = 1$, or equivalently, the complement of the union of all open arcs (or intervals, resp.) which have zero μ -measure. It is an elementary fact that every non-empty (relatively) open subset U of \mathbb{T} (or \mathbb{R} , resp.) can be written as a countable disjoint union of open arcs (or intervals, resp.): $U = \cup_{j=1}^n U_j$, where $n \in \mathbb{N} \cup \{\infty\}$ and U_j 's are exactly the connected components of U . Now, we give a characterisation of the set $\{\mu\}^c$ for every $\mu \in P_c(\mathbb{T})$.

Lemma 2.2. *Let $\mu \in P_c(\mathbb{T})$ and $\mathbb{T} \setminus S_\mu = \cup_{j=1}^n U_j$ where $n \in \mathbb{N} \cup \{\infty\}$ and U_j 's are the connected components of $\mathbb{T} \setminus S_\mu$. Then we have*

$$\{\mu\}^c = \bigcup_{j=1}^n \{\nu \in P_c(\mathbb{T}) : \nu(U_j) = 1\}. \quad (2.3)$$

In case when $\mu \in P_{ac}(\mathbb{T})$, then we have

$$\{\mu\}^{ac} = \bigcup_{j=1}^n \{\nu \in P_{ac}(\mathbb{T}) : \nu(U_j) = 1\}. \quad (2.4)$$

Moreover, the terms in the unions of (2.3) and (2.4) are exactly the connected components, i.e. the sets $\{\mu\}^c$ and $\{\mu\}^{ac}$ have exactly n connected components. In particular, $\{\mu\}^c$ (or $\{\mu\}^{ac}$) is non-empty and connected if and only if $S_\mu \neq \mathbb{T}$ and S_μ is connected; and $\{\mu\}^c = \emptyset$ (or $\{\mu\}^{ac} = \emptyset$) exactly when $S_\mu = \mathbb{T}$ holds.

Proof. We have

$$\begin{aligned} \{\mu\}^c &= \{\nu \in P_c(\mathbb{T}) : \exists A \in \mathcal{A}_0 \text{ such that} \\ &\text{either } \mu(A) = 0 \text{ and } \nu(A) = 1, \text{ or } \mu(\mathbb{T} \setminus A) = 0 \text{ and } \nu(\mathbb{T} \setminus A) = 1\}. \end{aligned}$$

Since we have $\mathbb{T} \setminus A \in \mathcal{A}_0$ for every $A \in \mathcal{A}_0$, and the measure of a closed arc and its interior are the same for continuous measures, we infer

$$\{\mu\}^c = \{\nu \in P_c(\mathbb{T}) : \exists A \in \mathcal{A}^{op} \text{ such that } \mu(A) = 0, \nu(A) = 1\}.$$

But $\mu(A) = 0$ holds if and only if $A \cap S_\mu = \emptyset$, therefore we conclude (2.3). The equation (2.4) can be obtained in a similar way.

For the other statement we make two observations. On one hand, if $\nu_1, \nu_2 \in \{\mu\}^c$, $\nu_1(U_{j_1}) = \nu_2(U_{j_2}) = 1$ and $j_1 \neq j_2$, then we easily conclude

$$1 \geq d(\nu_1, \nu_2) \geq |\nu_1(U_{j_1}) - \nu_2(U_{j_1})| = 1,$$

and thus $d(\nu_1, \nu_2) = 1$. This readily implies that the number of connected components of $\{\mu\}^c$ is at least n . On the other hand, if $\nu_1, \nu_2 \in \{\mu\}^c$, $\nu_1(U_{j_0}) = \nu_2(U_{j_0}) = 1$, then we consider

$$\gamma : [0, 1] \rightarrow P_c(\mathbb{T}), \quad \gamma(t) = (1-t) \cdot \nu_1 + t \cdot \nu_2.$$

Since we have

$$\begin{aligned} d(\gamma(s), \gamma(t)) &= \max\{|(1-t)\cdot\nu_1(A)+t\cdot\nu_2(A)-(1-s)\cdot\nu_1(A)-s\cdot\nu_2(A)|: A \subset \mathbb{T} \text{ is an arc}\} \\ &= |s-t| \cdot \max\{|\nu_1(A) - \nu_2(A)|: A \subset \mathbb{T} \text{ is an arc}\} = |s-t| \cdot d(\nu_1, \nu_2) \leq |s-t|, \end{aligned}$$

the curve γ connects ν_1 with ν_2 in $\{\nu \in P_c(\mathbb{T}): \nu(U_{j_0}) = 1\}$, whence the connectedness of $\{\nu \in P_c(\mathbb{T}): \nu(U_{j_0}) = 1\}$ is yielded. Finally, the proof for $\{\mu\}^{ac}$ is almost the same. \square

Next, we set

$$P_c^{cs}(\mathbb{T}) := \{\mu \in P_c(\mathbb{T}): \{\mu\}^c = \emptyset \text{ or } \{\mu\}^c \text{ is connected}\} = \{\mu \in P_c(\mathbb{T}): S_\mu \in \mathcal{A}^{cl}\}$$

and

$$P_{ac}^{cs}(\mathbb{T}) := \{\mu \in P_{ac}(\mathbb{T}): \{\mu\}^{ac} = \emptyset \text{ or } \{\mu\}^{ac} \text{ is connected}\} = \{\mu \in P_{ac}(\mathbb{T}): S_\mu \in \mathcal{A}^{cl}\}.$$

Let $\psi: P_c(\mathbb{T}) \rightarrow P_c(\mathbb{T})$ be an arbitrary surjective isometry with respect to the metric d . Since ψ is also a homeomorphism, we infer that

$$\{\mu\}^c = \emptyset \iff \psi(\{\mu\}^c) = \{\psi(\mu)\}^c = \emptyset$$

and

$$\{\mu\}^c \text{ is connected} \iff \{\psi(\mu)\}^c \text{ is connected},$$

whence

$$\psi(P_c^{cs}(\mathbb{T})) = P_c^{cs}(\mathbb{T})$$

is yielded. By (2.3) the following equivalences are straightforward:

$$\begin{aligned} S_\mu \subseteq S_\nu &\iff \{\nu\}^c \subseteq \{\mu\}^c \iff \{\psi(\nu)\}^c \subseteq \{\psi(\mu)\}^c \\ &\iff S_{\psi(\mu)} \subseteq S_{\psi(\nu)} \quad (\mu, \nu \in P_c(\mathbb{T})) \end{aligned} \quad (2.5)$$

and

$$\begin{aligned} S_\mu = S_\nu &\iff \{\nu\}^c = \{\mu\}^c \iff \{\psi(\nu)\}^c = \{\psi(\mu)\}^c \\ &\iff S_{\psi(\mu)} = S_{\psi(\nu)} \quad (\mu, \nu \in P_c(\mathbb{T})). \end{aligned} \quad (2.6)$$

Because of the above observations the following map can be defined:

$$\eta_\psi: \mathcal{A}^{cl} \rightarrow \mathcal{A}^{cl}, \quad \eta_\psi(S_\mu) = S_{\psi(\mu)} \quad (\mu \in P_c^{cs}(\mathbb{T})).$$

Since ψ^{-1} is also a surjective isometry, we obtain $\eta_\psi^{-1} = \eta_{\psi^{-1}}$. Clearly, we have the following property:

$$L \subseteq K \iff \eta_\psi(L) \subseteq \eta_\psi(K) \iff \eta_\psi^{-1}(L) \subseteq \eta_\psi^{-1}(K) \quad (L, K \in \mathcal{A}^{cl}). \quad (2.7)$$

In fact, more is true, which is proven in the next lemma. Let us point out that the verification of the analogues of the previous observations for $P_{ac}(\mathbb{T})$ is very similar. We will denote by λ the normalised arc-length measure on \mathbb{T} .

Lemma 2.3. *Let $\psi: P_c(\mathbb{T}) \rightarrow P_c(\mathbb{T})$ be a surjective isometry with respect to the metric d . Then there exists a homeomorphism $h: \mathbb{T} \rightarrow \mathbb{T}$ such that we have*

$$S_{\psi(\mu)} = \eta_\psi(S_\mu) = h(S_\mu) \quad (\mu \in P_c^{cs}(\mathbb{T})). \quad (2.8)$$

Moreover, if $\psi: P_{ac}(\mathbb{T}) \rightarrow P_{ac}(\mathbb{T})$ is a surjective isometry with respect to the metric d , then

$$S_{\psi(\mu)} = \eta_\psi(S_\mu) = h(S_\mu) \quad (\mu \in P_{ac}^{cs}(\mathbb{T})). \quad (2.9)$$

holds with a homeomorphism $h: \mathbb{T} \rightarrow \mathbb{T}$ which preserves sets with zero Lebesgue measure in both directions, i.e. we have

$$\lambda(A) = 0 \iff \lambda(h(A)) = 0 \quad (A \in \mathcal{B}_{\mathbb{T}}). \quad (2.10)$$

Proof. We begin with the first statement. Let us define a mapping $h: \mathbb{T} \rightarrow \mathbb{T}$ such that for each $t \in [-\pi, \pi)$ the point $h(\xi_t)$ is an arbitrary one lying in the intersection

$$\bigcap_{j=1}^{\infty} \eta_{\psi}(\{\xi_s: |s-t| \leq 1/j\}). \quad (2.11)$$

By (2.7) and completeness of the metric of \mathbb{T} , the above intersection is a non-empty closed arc, thus h is indeed a function. Let us assume for a moment that there is a $t \in [-\pi, \pi)$ such that the intersection in (2.11) is not a point, but a non-degenerate closed arc $K \in \mathcal{A}^{cl}$. In this case, (2.7) readily implies the contradiction $\#(\eta_{\psi}^{-1}(K)) \leq 1$. Therefore each of the above intersections has exactly one element, hence the map h is uniquely determined by the above properties.

Next, we observe that for every $t \in [-\pi, \pi)$ and $\{\alpha_k\}_{k=1}^{\infty}, \{\beta_k\}_{k=1}^{\infty} \subset [0, \infty)$ with $\alpha_k \searrow 0, \beta_k \searrow 0$ ($k \rightarrow \infty$), $\alpha_k + \beta_k > 0$ ($k \in \mathbb{N}$) we have

$$\bigcap_{k=1}^{\infty} \eta_{\psi}(\{\xi_s: s \in [t - \alpha_k, t + \beta_k]\}) = \{h(\xi_t)\}. \quad (2.12)$$

In order to verify this, we observe that there exists a non-decreasing sequence of positive integers $\{j_k\}_{k=1}^{\infty}$ such that $\lim_{k \rightarrow \infty} j_k = \infty$ and $[t - \alpha_k, t + \beta_k] \subseteq [t - 1/j_k, t + 1/j_k]$ ($k \in \mathbb{N}$). Thus, again by (2.7), the intersection in (2.12) is a subset of $\{h(\xi_t)\}$. But clearly this intersection cannot be empty.

The above observations imply $h(K) \subseteq \eta_{\psi}(K)$ ($K \in \mathcal{A}^{cl}$). Let $K \in \mathcal{A}^{cl}$ be arbitrary, and let us consider a point $\chi \in \eta_{\psi}(K)$. There is a monotone decreasing sequence $\{K_j\}_{j=1}^{\infty} \subset \mathcal{A}^{cl}$ such that $K_j \subset \eta_{\psi}(K)$ ($j \in \mathbb{N}$) and $\bigcap_{j=1}^{\infty} K_j = \{\chi\}$. Clearly, $\{\eta_{\psi}^{-1}(K_j)\}_{j=1}^{\infty} = \{\eta_{\psi}^{-1}(K_j)\}_{j=1}^{\infty} \subset \mathcal{A}^{cl}$ is a monotone decreasing sequence of closed arcs such that $\eta_{\psi}^{-1}(K_j) \subset K$ and $\#(\bigcap_{j=1}^{\infty} \eta_{\psi}^{-1}(K_j)) = \#(\bigcap_{j=1}^{\infty} \eta_{\psi}^{-1}(K_j)) = 1$ ($j \in \mathbb{N}$). Therefore $h(K) = \eta_{\psi}(K)$, which verifies (2.8).

It remains to show that h is a homeomorphism. Let $\mu, \nu \in P_c^{cs}(\mathbb{T})$ be arbitrary and set $K = S_{\mu}, L = S_{\nu}$. We have

$$\{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T}) = \{\vartheta \in P_c^{cs}(\mathbb{T}): \vartheta(\mathbb{T} \setminus (K \cup L)) = 1\}.$$

On one hand, if $K \cap L = \emptyset$, then $\mathbb{T} \setminus (K \cup L)$ has exactly two connected components U_1 and U_2 , which implies that in this case $\{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T}) = \bigcup_{j=1}^2 \{\vartheta \in P_c^{cs}(\mathbb{T}): \vartheta(U_j) = 1\}$. But for any choices $\vartheta_j \in \{\vartheta \in P_c^{cs}(\mathbb{T}): \vartheta(U_j) = 1\}$ ($j = 1, 2$), we clearly have $d(\vartheta_1, \vartheta_2) = 1$, which implies that $\{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T})$ is not a connected set in this case. On the other hand, if $K \cap L \neq \emptyset$, then $\mathbb{T} \setminus (K \cup L)$ is a connected open set, moreover, we claim that $\{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T})$ is connected. Let $\vartheta_1, \vartheta_2 \in \{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T})$. We choose a $\tilde{\vartheta} \in P_c^{cs}(\mathbb{T})$ with

$$S_{\vartheta_1} \cup S_{\vartheta_2} \subseteq S_{\tilde{\vartheta}} \subseteq (\mathbb{T} \setminus (K \cup L))^{-},$$

where $\bar{\cdot}$ means the closure of a given set, and consider the following path

$$\gamma: [0, 1] \rightarrow P_c(\mathbb{T}), \quad \gamma(t) = (t - t^2) \cdot \tilde{\vartheta} + (1 - t + t^2) \cdot [(1 - t) \cdot \vartheta_1 + t \cdot \vartheta_2].$$

Clearly, $\gamma([0, 1]) \subseteq \{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T})$, $\gamma(0) = \vartheta_1$, $\gamma(1) = \vartheta_2$, and a straightforward computation gives the continuity of γ . Thus, indeed, the set $\{\mu\}^c \cap \{\nu\}^c \cap P_c^{cs}(\mathbb{T})$ is connected.

Applying these observations we get the following:

$$K \cap L = \emptyset \iff h(K) \cap h(L) = \emptyset \quad (K, L \in \mathcal{A}^{cl}). \quad (2.13)$$

Clearly, $h(\mathbb{T}) = \eta_\psi(\mathbb{T}) = \mathbb{T}$, thus h is surjective. Let us consider two different points ξ_s and ξ_t ($s \neq t, s, t \in [-\pi, \pi)$). Then there are two disjoint closed arcs $K_s, K_t \in \mathcal{A}^{cl}$ such that $\xi_s \in K_s$ and $\xi_t \in K_t$. By (2.13) we have $h(K_s) \cap h(K_t) = \emptyset$, whence we infer the bijectivity of h . Finally, let $\{t_n\}_{n=1}^\infty \subset \mathbb{R}$ be a sequence such that $t_n \searrow t$ or $t_n \nearrow t$ ($n \rightarrow \infty$) and $|t_1 - t| < \frac{\pi}{2}$, and let K_n be the shorter closed arc with endpoints ξ_t and ξ_{t_n} . Since we have $\{h(\xi_t)\} = \bigcap_{n=1}^\infty h(K_n)$ and $h(\xi_{t_n}) \in h(K_n)$ ($n \in \mathbb{N}$), the continuity of h is yielded. But h is a bijective continuous map of the compact Hausdorff space \mathbb{T} , thus we conclude that h is a homeomorphism. This completes the proof for the first case.

For the second case, we obtain in a similar way as above that there is a homeomorphism $h: \mathbb{T} \rightarrow \mathbb{T}$ such that (2.9) is satisfied. Let $K \subseteq \mathbb{T}$ be a compact set with $\lambda(K) > 0$, and $\mathbb{T} \setminus K = \bigcup_{j=1}^n U_j$ ($n \in \mathbb{N} \cup \{\infty\}$) where the union is disjoint and $U_j \in \mathcal{A}^{op}$ for every j . We intend to show that $\lambda(h(K)) > 0$ is satisfied. Obviously, there exists a $\mu \in P_{ac}(\mathbb{T})$ with $S_\mu = K$. By (2.5) and (2.9) we infer

$$h(K) = h\left(\bigcap_{j=1}^n (\mathbb{T} \setminus U_j)\right) = \bigcap_{j=1}^n h(\mathbb{T} \setminus U_j) = \bigcap_{j=1}^n \eta_\psi(\mathbb{T} \setminus U_j) \supseteq S_{\psi(\mu)}.$$

Thus we obtain $0 < \lambda(S_{\psi(\mu)}) \leq \lambda(h(K))$. Next, if $A \in \mathcal{B}_{\mathbb{T}}$ has positive Lebesgue measure, then by regularity we infer that there is a compact subset $K \subseteq A$ such that K has still positive Lebesgue measure. Therefore we conclude

$$\lambda(h(A)) = 0 \implies \lambda(A) = 0 \quad (A \in \mathcal{B}_{\mathbb{T}}).$$

For the reverse direction we only have to observe the following property which we can conclude from (2.9):

$$S_{\psi^{-1}(\mu)} = \eta_{\psi^{-1}}(S_\mu) = \eta_\psi^{-1}(S_\mu) = h^{-1}(S_\mu) \quad (\mu \in P_{ac}^{cs}(\mathbb{T})).$$

This completes the proof. \square

We proceed with the verification of the next lemma where measures on \mathbb{R} are considered.

Lemma 2.4. *Let $I \in \mathcal{I}$ be a closed (possibly unbounded) interval, and $\mu \in P_c(\mathbb{R})$. We have the following equation:*

$$1 - \mu(I) = \inf\{d_{Ku}(\mu, \vartheta) : \vartheta \in P_c(\mathbb{R}), S_\vartheta \subseteq I\}. \quad (2.14)$$

Furthermore, if $\mu \in P_{ac}(\mathbb{R})$, then we have

$$1 - \mu(I) = \inf\{d_{Ku}(\mu, \vartheta) : \vartheta \in P_{ac}(\mathbb{R}), S_\vartheta \subseteq I\}. \quad (2.15)$$

Proof. We will only deal with (2.14), since (2.15) can be handled in a similar way. We have

$$d_{Ku}(\mu, \vartheta) \geq |\mu(I) - \vartheta(I)| = 1 - \mu(I) \quad (\vartheta \in P_c(\mathbb{R}), S_\vartheta \subseteq I),$$

which implies $1 - \mu(I) \leq \inf\{d_{Ku}(\mu, \vartheta) : \vartheta \in P_c(\mathbb{R}), S_\vartheta \subseteq I\}$. If $\mu(I) = 0$, then we immediately obtain (2.14), thus in the sequel we may assume that we have $\mu(I) > 0$.

Let us define $\nu \in P(\mathbb{R})$ by

$$\nu(A) = \frac{\mu(A \cap I)}{\mu(I)} \quad (A \in \mathcal{B}_{\mathbb{R}}).$$

Clearly, we have $\nu \in P_c(\mathbb{R})$ and $S_\nu \subseteq I$, moreover

$$|\nu(J) - \mu(J)| = \nu(J) - \mu(J) = \frac{\mu(J)}{\mu(I)}(1 - \mu(I)) \leq 1 - \mu(I) \quad (J \in \mathcal{I}, J \subseteq I). \quad (2.16)$$

Now, we consider an arbitrary interval $\tilde{I} \in \mathcal{I}$. There are two possibilities: either $\mu(\tilde{I}) - \nu(\tilde{I}) \leq 0$, or $\mu(\tilde{I}) - \nu(\tilde{I}) > 0$. In case of the first one, we have the following estimation:

$$\left| \mu(\tilde{I}) - \nu(\tilde{I}) \right| = \nu(\tilde{I}) - \mu(\tilde{I}) \leq \nu(I \cap \tilde{I}) - \mu(I \cap \tilde{I}) \leq 1 - \mu(I),$$

where we used (2.16). For the second possibility, $\tilde{I} \subseteq I$ is impossible, thus, again by (2.16), we estimate in the following way:

$$\begin{aligned} \left| \mu(\tilde{I}) - \nu(\tilde{I}) \right| &= \mu(\tilde{I}) - \nu(\tilde{I}) = \nu(\mathbb{R} \setminus \tilde{I}) - \mu(\mathbb{R} \setminus \tilde{I}) = \nu((\mathbb{R} \setminus \tilde{I}) \cap I) - \mu(\mathbb{R} \setminus \tilde{I}) \\ &\leq \nu((\mathbb{R} \setminus \tilde{I}) \cap I) - \mu((\mathbb{R} \setminus \tilde{I}) \cap I) \leq 1 - \mu(I), \end{aligned}$$

where we observed that $(\mathbb{R} \setminus \tilde{I}) \cap I$ is an interval. Since $1 - \mu(I) = \nu(I) - \mu(I)$, we conclude $d_{Ku}(\mu, \nu) = 1 - \mu(I)$. Therefore (2.14) follows, which ends the proof. \square

For an arbitrary $\mu \in P_c(\mathbb{R})$ we define the measure $\mu \circ \tau^{-1}$ by $(\mu \circ \tau^{-1})(A) = \mu(\tau^{-1}(A))$ ($A \in \mathcal{B}_{\mathbb{T}}$), where $\tau^{-1}(A)$ is the inverse image of A . It is straightforward that the transformation $\mu \mapsto \mu \circ \tau^{-1}$ is a bijection between $P_c(\mathbb{R})$ and $P_c(\mathbb{T})$. Furthermore, μ is absolutely continuous if and only if $\mu \circ \tau^{-1} \in P_{ac}(\mathbb{T})$. We define in a very similar way the measure $\tilde{\mu} \circ h \in P_c(\mathbb{T})$ where $\tilde{\mu} \in P_c(\mathbb{T})$ and $h: \mathbb{T} \rightarrow \mathbb{T}$ is a homeomorphism. Let us note that in case when $\tilde{\mu} \in P_{ac}(\mathbb{T})$, then $\tilde{\mu} \circ h$ is not necessarily absolutely continuous.

Now, we are in the position to prove our first theorem.

Proof of Theorem 1.1. Let

$$\psi: P_c(\mathbb{T}) \rightarrow P_c(\mathbb{T}), \quad \psi(\tilde{\mu}) = \phi(\tilde{\mu} \circ \tau) \circ \tau^{-1},$$

It is clear that ϕ is a surjective isometry with respect to the Kuiper metric if and only if ψ is a surjective isometry with respect to the metric d .

By Lemma 2.3, there exists a homeomorphism $h: \mathbb{T} \rightarrow \mathbb{T}$ such that

$$S_{\psi(\tilde{\mu})} = h(S_{\tilde{\mu}}) \quad (\tilde{\mu} \in P_c^{cs}(\mathbb{T})).$$

Let us consider the following mapping:

$$\psi_1: P_c(\mathbb{T}) \rightarrow P_c(\mathbb{T}), \quad \psi_1(\tilde{\mu}) = (\psi(\tilde{\mu})) \circ h.$$

It is straightforward to see that ψ_1 is a surjective isometry if and only if ψ is, and that we have

$$S_{\psi_1(\tilde{\mu})} = h^{-1}(S_{\psi(\tilde{\mu})}) = S_{\tilde{\mu}} \quad (\tilde{\mu} \in P_c^{cs}(\mathbb{T})).$$

We define the transformation

$$\phi_1: P_c(\mathbb{R}) \rightarrow P_c(\mathbb{R}), \quad \phi_1(\mu) = \psi_1(\mu \circ \tau^{-1}) \circ \tau.$$

It is apparent that ϕ_1 is a surjective isometry if and only if ϕ is, and that we have

$$S_{\phi_1(\mu)} = S_\mu \quad (\mu \in P_c^{cs}(\mathbb{R})), \quad (2.17)$$

where

$$P_c^{cs}(\mathbb{R}) = \{\mu \in P_c(\mathbb{R}): S_\mu \in \mathcal{I} \text{ or } \mathbb{R} \setminus S_\mu \in \mathcal{I}\}.$$

The following equivalence follows easily from the definition of ϕ_1 and the property (2.5) for ψ_1 :

$$S_\mu \subseteq S_\nu \iff S_{\phi_1(\mu)} \subseteq S_{\phi_1(\nu)} \quad (\mu, \nu \in P_c(\mathbb{R})). \quad (2.18)$$

Now, let μ be an arbitrary continuous Borel probability measure on \mathbb{R} . Since we have $\psi_1(P_c^{cs}(\mathbb{T})) = P_c^{cs}(\mathbb{T})$, we infer $\phi_1(P_c^{cs}(\mathbb{R})) = P_c^{cs}(\mathbb{R})$. This, (2.17), (2.18) and Lemma 2.4 implies the following for every closed interval $I \in \mathcal{I}$:

$$\begin{aligned} \phi_1(\mu)(I) &= 1 - \inf\{d_{Ku}(\phi_1(\mu), \vartheta) : \vartheta \in P_c(\mathbb{R}), S_\vartheta \subseteq I\} \\ &= 1 - \inf\{d_{Ku}(\mu, \phi_1^{-1}(\vartheta)) : \vartheta \in P_c(\mathbb{R}), S_\vartheta \subseteq I\} \\ &= 1 - \inf\{d_{Ku}(\mu, \vartheta) : \vartheta \in P_c(\mathbb{R}), S_\vartheta \subseteq I\} = \mu(I). \end{aligned}$$

But this immediately implies $\phi_1(\mu) = \mu$, and thus that ϕ_1 is the identity map. It is tedious, but straightforward, to check that transforming back to our original map ϕ yields (1.2) where

$$x = \begin{cases} \infty & \text{if } h(-1) = -1 \\ \tau^{-1}(h(-1)) & \text{otherwise} \end{cases}$$

and g is the continuous extension of $\tau^{-1} \circ h^{-1} \circ \tau \circ r_x^{-1}$. Note that the latter function is not defined in at most two points of \mathbb{R} (depending on the actual value of x and the function h). However, if we consider r_x as a bijective function of the one-point compactification of \mathbb{R} (which is topologically equivalent to \mathbb{T}), then it is not hard to see that g (as the continuous extension of $\tau^{-1} \circ h^{-1} \circ \tau \circ r_x^{-1}$) makes sense and that it is indeed a homeomorphism of \mathbb{R} . \square

The proof of our second result is quite similar to the above one, therefore we only present its sketch.

Proof of Theorem 1.2. The definitions of ψ , ψ_1 and ϕ_1 are similar to the above definitions. It is straightforward that each of these transformations is a surjective Kuiper isometry of $P_{ac}(\mathbb{T})$ if and only if ϕ is. While transforming back the identity map (i.e. ϕ_1) to ϕ we observe that the homeomorphism g from the statement has the property

$$m(B) = 0 \iff m(g(B)) = 0 \quad (B \in \mathcal{B}_{\mathbb{R}}).$$

Since every homeomorphism of \mathbb{R} is either monotone increasing or decreasing, we infer that g and g^{-1} are of bounded variation on every compact interval. Therefore the famous Banach–Zarecki theorem (see e.g. [3, 5]) implies that they are both locally absolutely continuous functions. This completes one direction of the statement. The other direction is a rather easy calculation. \square

3. PROOF IN THE GENERAL CASE

In order to prove our last theorem, we need to introduce a new type of support for Borel probability measures. Let $\mu \in P(\mathbb{R})$ be an arbitrary measure, and let us define the following sets

$$N_\mu = \cup\{I \in \mathcal{I} : \mu(I) = 0\}, \quad C_\mu = \mathbb{R} \setminus N_\mu.$$

The set C_μ will be called the co-interval support of μ . Clearly, this is the unique smallest set such that its complement is a union of non-degenerate intervals and μ is concentrated on it. This is an analogue of the usual closed support S_μ , however, as we shall see this notion is more useful here. Recall that S_μ is the complement of

the union of open intervals with zero μ -measure. Therefore the following properties of the co-interval support are straightforward:

$$C_\mu \subseteq S_\mu, \quad \overline{C_\mu} = S_\mu \quad \text{and} \quad \#(S_\mu \setminus C_\mu) \leq \aleph_0.$$

If $\mathcal{M} \subseteq P(\mathbb{R})$, then let

$$\mathcal{M}^1 = \{\nu \in P(\mathbb{R}) : d_{Ku}(\mu, \nu) = 1 \quad \forall \mu \in \mathcal{M}\}.$$

We call a probability measure which is concentrated on a point $x \in \mathbb{R}$ a Dirac measure, and we will denote it by δ_x .

We begin with the following metric characterisation of Dirac measures.

Lemma 3.1. *For an arbitrary $\mu \in P(\mathbb{R})$ we have $\#((\{\mu\}^1)^1) = 1$ if and only if μ is a Dirac measure.*

Proof. Let $\nu \in P(\mathbb{R})$ be an arbitrary measure which is absolutely continuous with respect to μ . We show that $\nu \in (\{\mu\}^1)^1$. By Lemma 2.1, for any $\vartheta \in \{\mu\}^1$ there exists an $I \in \mathcal{I}_0$ such that either we have $\mu(I) = 0$ and $\vartheta(I) = 1$, or $\mu(\mathbb{R} \setminus I) = 0$ and $\vartheta(\mathbb{R} \setminus I) = 1$. Since ν is absolutely continuous with respect to μ , we immediately infer either $\nu(I) = 0$ and $\vartheta(I) = 1$, or $\nu(\mathbb{R} \setminus I) = 0$ and $\vartheta(\mathbb{R} \setminus I) = 1$, which implies $d_{Ku}(\vartheta, \nu) = 1$. Since this holds for every $\vartheta \in \{\mu\}^1$, the relation $\nu \in (\{\mu\}^1)^1$ follows.

Now, if μ is not a Dirac measure, then it is straightforward that there are infinitely many $\nu \in P(\mathbb{R})$ which is absolutely continuous with respect to μ . Hence, we obtain that $\#((\{\mu\}^1)^1) = 1$ implies that μ is a Dirac measure.

Finally, we show that $(\{\delta_x\}^1)^1 = \{\delta_x\}$. Let $\delta_x \in P(\mathbb{R})$ be a Dirac measure with $x \in \mathbb{R}$. By Lemma 2.1, we have $\vartheta \in \{\delta_x\}^1$ if and only if there exists a possibly degenerate interval $I \in \mathcal{I}_0$ such that either $\delta_x(I) = 0$ and $\vartheta(I) = 1$, or $\delta_x(\mathbb{R} \setminus I) = 0$ and $\vartheta(\mathbb{R} \setminus I) = 1$. Obviously, in the first case we have $x \in \mathbb{R} \setminus I$, and in the second one $x \in I$. Therefore we infer the equation

$$\{\delta_x\}^1 = \{\vartheta \in P(\mathbb{R}) : \vartheta(\{x\}) = 0\}.$$

Assume that $\nu \in (\{\delta_x\}^1)^1$, i.e. $d_{Ku}(\nu, \vartheta) = 1$ for every $\vartheta \in P(\mathbb{R})$, $\vartheta(\{x\}) = 0$. Let us suppose for a moment that $\nu(\mathbb{R} \setminus \{x\}) > 0$, and let ϑ be defined by

$$\vartheta(B) = \frac{1}{\nu(\mathbb{R} \setminus \{x\})} \cdot \nu(B \setminus \{x\}) \quad (B \in \mathcal{B}_{\mathbb{R}}).$$

It is clear that $\vartheta \in \{\delta_x\}^1$ and that ϑ is absolutely continuous with respect to ν . Thus $\nu(I) = 0$ and $\vartheta(I) = 1$ cannot be satisfied simultaneously when $I \in \mathcal{I}_0$, and the same holds for $\nu(\mathbb{R} \setminus I) = 0$ and $\vartheta(\mathbb{R} \setminus I) = 1$. This implies $d_{Ku}(\nu, \vartheta) < 1$, a contradiction. Hence $\nu(\mathbb{R} \setminus \{x\}) = 0$ follows, and we obtain $(\{\delta_x\}^1)^1 = \{\delta_x\}$, which completes the proof. \square

We proceed with verifying the following property of surjective Kuiper isometries.

Lemma 3.2. *Let $\phi: P(\mathbb{R}) \rightarrow P(\mathbb{R})$ be a surjective isometry with respect to the Kuiper distance. Then there exists a bijection $f: \mathbb{R} \rightarrow \mathbb{R}$ such that*

$$\phi(\mu)(\{f(x)\}) = \mu(\{x\}) \quad (\mu \in P(\mathbb{R}), x \in \mathbb{R}).$$

Proof. Let Δ denote the set of all Dirac measures of $P(\mathbb{R})$. Since $\phi((\{\mu\}^1)^1) = (\phi(\{\mu\}^1))^1 = (\{\phi(\mu)\}^1)^1$ holds, by Lemma 3.1 we obtain $\phi(\Delta) = \Delta$, i.e. we have a bijection $f: \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\phi(\delta_x) = \delta_{f(x)} \quad (x \in \mathbb{R}).$$

Let $\mu \in P(\mathbb{R})$ be an arbitrary measure. We state that

$$d_{Ku}(\mu, \delta_x) = 1 - \mu(\{x\}) \quad (x \in \mathbb{R}). \quad (3.1)$$

On one hand, since $\{x\} \in \mathcal{I}_0$ and $|\delta_x(\{x\}) - \mu(\{x\})| = 1 - \mu(\{x\})$, we obtain $d_{Ku}(\mu, \delta_x) \geq 1 - \mu(\{x\})$. On the other hand, let $I \in \mathcal{I}_0$ be an arbitrary, possibly degenerate, interval. If $x \in I$, then we have $|\delta_x(I) - \mu(I)| = 1 - \mu(I) \leq 1 - \mu(\{x\})$; and if $x \notin I$, then we have $|\delta_x(I) - \mu(I)| = \mu(I) \leq \mu(\mathbb{R} \setminus \{x\}) = 1 - \mu(\{x\})$. Therefore we conclude (3.1).

Finally, observing the following for every $x \in \mathbb{R}$ and $\mu \in P(\mathbb{R})$ completes the proof:

$$\mu(\{x\}) = 1 - d_{Ku}(\mu, \delta_x) = 1 - d_{Ku}(\phi(\mu), \phi(\delta_x)) = 1 - d_{Ku}(\phi(\mu), \delta_{f(x)}) = \phi(\mu)(\{f(x)\}).$$

□

It is straightforward that for every $\mu \in P(\mathbb{R}) \setminus \Delta$ we can write N_μ as a disjoint union $N_\mu = (\mathbb{R} \setminus \text{conv}(C_\mu)) \cup (\cup_{j=1}^n I_j)$ where $n \in \mathbb{N} \cup \{0, \infty\}$, $I_j \in \mathcal{I}$ for every j , and $\text{conv}(\cdot)$ denotes the convex hull of a given set. In fact, the I_j 's are the bounded connected components of N_μ , and in case when there is at least one unbounded connected component, then the union of these components is exactly $\mathbb{R} \setminus \text{conv}(C_\mu)$. We note that the intersection of the closure of two connected components can be empty or one point.

Now, we are ready to prove an analogue of Lemma 2.2, where we describe $\{\mu\}^1 \setminus \Delta$ instead of $\{\mu\}^1$. The reason for this is that the presence of Dirac measures would cause some technical problems.

Lemma 3.3. *Let $\mu \in P(\mathbb{R}) \setminus \Delta$ and let us write $N_\mu = (\mathbb{R} \setminus \text{conv}(C_\mu)) \cup (\cup_{j=1}^n I_j)$ where the right-hand side is the above mentioned disjoint union. Then we have*

$$\{\mu\}^1 \setminus \Delta = \left\{ \nu \in P(\mathbb{R}) \setminus \Delta : \nu(\mathbb{R} \setminus \text{conv}(C_\mu)) = 1 \right\} \bigcup \left(\bigcup_{j=1}^n \left\{ \nu \in P(\mathbb{R}) \setminus \Delta : \nu(I_j) = 1 \right\} \right). \quad (3.2)$$

Proof. By Lemma 2.1 we have

$$\begin{aligned} \{\mu\}^1 \setminus \Delta &= \{ \nu \in P(\mathbb{R}) \setminus \Delta : \exists I \in \mathcal{I}_0 \nu(I) = 0 \text{ and } \mu(I) = 1 \} \\ &\cup \{ \nu \in P(\mathbb{R}) \setminus \Delta : \exists I \in \mathcal{I}_0 \nu(I) = 1 \text{ and } \mu(I) = 0 \} \\ &= \{ \nu \in P(\mathbb{R}) \setminus \Delta : \exists I \in \mathcal{I} \nu(I) = 0 \text{ and } \mu(I) = 1 \} \\ &\cup \{ \nu \in P(\mathbb{R}) \setminus \Delta : \exists I \in \mathcal{I} \nu(I) = 1 \text{ and } \mu(I) = 0 \}. \end{aligned}$$

On one hand, if we have $\nu(I) = 0$ and $\mu(I) = 1$ with some $I \in \mathcal{I}$, then we automatically infer $\text{conv}(C_\mu) \subseteq I$. Therefore this case is equivalent to $\nu(\mathbb{R} \setminus \text{conv}(C_\mu)) = 1$. On the other hand, if we have $\mu(I) = 0$ and $\nu(I) = 1$ with some $I \in \mathcal{I}$, then $I \subseteq N_\mu$ is yielded, or equivalently $I \subseteq I_j$ with some j or $I \subseteq \mathbb{R} \setminus \text{conv}(C_\mu)$. Therefore this case is equivalent to either $\nu(I_j) = 1$ for some j , or that the ν -measure of one of the unbounded components of N_μ is 1. Noting that this latter possibility implies $\nu(\mathbb{R} \setminus \text{conv}(C_\mu)) = 1$ ends our proof. □

We continue with the following lemma.

Lemma 3.4. *The function f defined in Lemma 3.2 is a homeomorphism (i.e. a monoton bijection).*

Proof. First, we show the following property:

$$C_\mu \subseteq C_\nu \iff C_{\phi(\mu)} \subseteq C_{\phi(\nu)} \quad (\mu, \nu \in P(\mathbb{R})). \quad (3.3)$$

If $\mu, \nu \notin \Delta$, then by Lemma 3.3 the following equivalence is yielded:

$$C_\mu \subseteq C_\nu \iff (\{\nu\}^1)^1 \setminus \Delta \subseteq (\{\mu\}^1)^1 \setminus \Delta \quad (\mu, \nu \in P(\mathbb{R}) \setminus \Delta).$$

Since we have $\phi(P(\mathbb{R}) \setminus \Delta) = P(\mathbb{R}) \setminus \Delta$, we get $\phi(\mu), \phi(\nu) \notin \Delta$. Obviously, we have $\phi((\{\mu\}^1)^1 \setminus \Delta) = \phi((\{\mu\}^1)^1) \setminus \Delta = (\{\phi(\mu)\}^1)^1 \setminus \Delta$, whence we conclude (3.3) in this case. Next, we suppose that $\nu = \delta_x$ holds with some $x \in \mathbb{R}$. Clearly, $C_\mu \subseteq C_{\delta_x}$ is equivalent to $\mu = \delta_x$, which holds if and only if $\phi(\mu) = \phi(\delta_x) = \delta_{f(x)}$. This latter equation is valid exactly when we have $C_{\phi(\mu)} \subseteq C_{\phi(\delta_x)}$, which gives us (3.3) in this case. Finally, let us assume that we have $\nu \notin \Delta$, $\mu = \delta_x$ and $C_{\delta_x} \subseteq C_\nu$ with some $x \in \mathbb{R}$. We consider a point $y \in C_\nu$, $x \neq y$. Since we have $C_{\frac{1}{2}\delta_x + \frac{1}{2}\delta_y} \subseteq C_\nu$, by the previous cases, we obtain $C_{\phi(\frac{1}{2}\delta_x + \frac{1}{2}\delta_y)} \subseteq C_{\phi(\nu)}$. Moreover, by Lemma 3.2, we have $\phi(\frac{1}{2}\delta_x + \frac{1}{2}\delta_y) = \frac{1}{2}\phi(\delta_x) + \frac{1}{2}\phi(\delta_y) = \frac{1}{2}\delta_{f(x)} + \frac{1}{2}\delta_{f(y)}$, thus we obtain $C_{\phi(\delta_x)} = C_{\delta_{f(x)}} \subseteq C_{\phi(\frac{1}{2}\delta_x + \frac{1}{2}\delta_y)} \subseteq C_{\phi(\nu)}$. This verifies (3.3) in this case in one direction. The other direction follows from the fact that ϕ^{-1} is also a surjective Kuiper isometry. Therefore we conclude (3.3) in general.

Next, by Lemma 3.2, we have $\phi(P_c(\mathbb{R})) = P_c(\mathbb{R})$. Theorem 1.1 gives us a homeomorphism $g: \mathbb{R} \rightarrow \mathbb{R}$ and an $x \in \mathbb{R} \cup \{\infty\}$ such that we have

$$\phi(\mu) = \mu \circ (g \circ r_x) \quad (\mu \in P_c(\mathbb{R})),$$

and consequently we obtain

$$S_{\phi(\mu)} = \begin{cases} g^{-1}(S_\mu) & \text{if } x = \infty \\ r_0 \circ g^{-1}(S_\mu) + x & \text{if } x \in \mathbb{R} \text{ and } S_\mu \text{ is compact} \\ \{x\} \cup (r_0 \circ g^{-1}(S_\mu) + x) & \text{if } x \in \mathbb{R} \text{ and } S_\mu \text{ is not compact} \end{cases} \quad (\mu \in P_c(\mathbb{R})).$$

Let us suppose for a moment that we have $x \in \mathbb{R}$, and consider a sequence of continuous measures $\{\mu_n\}_{n=1}^\infty$ with

$$C_{\mu_n} = \begin{cases} \left(g\left(-\frac{1}{n}\right), g\left(\frac{1}{n}\right)\right) & \text{if } g \text{ is monotone increasing} \\ \left(g\left(\frac{1}{n}\right), g\left(-\frac{1}{n}\right)\right) & \text{if } g \text{ is monotone decreasing} \end{cases}.$$

Since $\{g(0)\} = C_{\delta_{g(0)}} \subseteq C_{\mu_n}$, we infer $\{f(g(0))\} = C_{\phi(\delta_{g(0)})} \subseteq C_{\phi(\mu_n)} = \mathbb{R} \setminus [-n+x, n+x]$ for every $n \in \mathbb{N}$, which gives the contradiction $\{f(g(0))\} \subseteq \bigcap_{n=1}^\infty (\mathbb{R} \setminus [-n+x, n+x]) = \emptyset$. Therefore we conclude that

$$\phi(\mu) = \mu \circ g \text{ and } S_{\phi(\mu)} = g^{-1}(S_\mu) \quad (\mu \in P_c(\mathbb{R})),$$

holds with some homeomorphism g .

Now, we consider a number $t \in \mathbb{R}$ and a sequence $\{\nu_n\}_{n=1}^\infty \subset P_c(\mathbb{R})$ such that

$$C_{\nu_n} = \begin{cases} \left(g\left(g^{-1}(t) - \frac{1}{n}\right), g\left(g^{-1}(t) + \frac{1}{n}\right)\right) & \text{if } g \text{ is monotone increasing} \\ \left(g\left(g^{-1}(t) + \frac{1}{n}\right), g\left(g^{-1}(t) - \frac{1}{n}\right)\right) & \text{if } g \text{ is monotone decreasing} \end{cases}.$$

Since we have $\{t\} = C_{\delta_t} \subseteq C_{\nu_n}$ ($n \in \mathbb{N}$), we obtain $\{f(t)\} = C_{\phi(\delta_t)} \subseteq \bigcap_{n=1}^\infty C_{\phi(\nu_n)} = \{g^{-1}(t)\}$, which implies $g = f^{-1}$. As a consequence we have that f is a homeomorphism, which makes the proof complete. \square

Before we prove Theorem 1.3, we need one further statement.

Lemma 3.5. *The set of all purely atomic probability measures are dense in $P(\mathbb{R})$ with respect to the Kuiper distance.*

Proof. Here it is easier to consider distribution functions instead of measures. Let us observe that on the space of all distribution functions the Kuiper distance is equivalent to the Kolmogorov–Smirnov distance, which is obtained from the supremum norm. Since the statement is quite straightforward to show in the Kolmogorov–Smirnov distance, we also have it for the Kuiper distance. \square

Finally, we are in the position to present the verification of our last result.

Proof of Theorem 1.3. Assume that $\phi: P(\mathbb{R}) \rightarrow P(\mathbb{R})$ is a surjective Kuiper isometry. Let us consider the function $f: \mathbb{R} \rightarrow \mathbb{R}$ which was defined in Lemma 3.2 and which is a homeomorphism by Lemma 3.4. Therefore

$$\phi_1: P(\mathbb{R}) \rightarrow P(\mathbb{R}), \quad \phi_1(\mu) = \phi(\mu) \circ f \quad (\mu \in P(\mathbb{R}))$$

is also a surjective isometry with respect to the Kuiper distance, and we have

$$(\phi_1(\mu))(\{x\}) = \mu(\{x\}) \quad (x \in \mathbb{R}, \mu \in P(\mathbb{R})).$$

Consequently, we obtain

$$\phi_1 \left(\sum_{j=1}^n \alpha_j \delta_{x_j} \right) = \sum_{j=1}^n \alpha_j \delta_{x_j} \quad \left(j \in \mathbb{N} \cup \{\infty\}, x_j \in \mathbb{R}, \alpha_j > 0, \sum_{j=1}^n \alpha_j = 1 \right).$$

Since isometries are automatically continuous, Lemma 3.5 implies that ϕ_1 has to be the identity map. Therefore, we conclude $\phi(\mu) = \mu \circ f^{-1}$ ($\mu \in P(\mathbb{R})$), which completes the proof. \square

ACKNOWLEDGMENTS

The author was also supported by the "Lendület" Program (LP2012-46/2012) of the Hungarian Academy of Sciences and by the Hungarian National Research, Development and Innovation Office – NKFIH (grant no. K115383).

REFERENCES

1. L. Davies, *Data analysis and approximate models, Model choice, location-scale, analysis of variance, nonparametric regression and image analysis*, Monographs on Statistics and Applied Probability, 133. CRC Press, Boca Raton, FL, 2014.
2. P.L. Davies, and A. Kovac, Densities, spectral densities and modality, *Ann. Statist.* **32** (2004), 1093–1136.
3. J. Duda and L. Zajíček, The Banach–Zarecki theorem for functions with values in metric spaces, *Proc. Amer. Math. Soc.* **133** (2005), 3631–3633.
4. G. Dolinar and L. Molnár, Isometries of the space of distribution functions with respect to the Kolmogorov–Smirnov metric, *J. Math. Anal. Appl.* **348** (2008), 494–498.
5. V. Ene, An elementary proof of the Banach–Zarecki theorem, *Real Anal. Exchange* **23** (1997/98), 295–301.
6. N.H. Kuiper, On a metric in the space of random variables, *Statistica Neerlandica* **16** (1962) 231–235.
7. L. Molnár, Kolmogorov–Smirnov isometries and affine automorphisms of spaces of distribution functions, *Cent. Eur. J. Math.* **9** (2011), 789–796.
8. L. Molnár, Lévy isometries of the space of probability distribution functions, *J. Math. Anal. Appl.* **380** (2011), 847–852.
9. L. Molnár and P. Szokol, Kolmogorov–Smirnov isometries of the space of generalized distribution functions, *Math. Slovaca* **64** (2014), 433–444.

MTA-SZTE ANALYSIS AND STOCHASTICS RESEARCH GROUP, BOLYAI INSTITUTE, UNIVERSITY OF SZEGED, H-6720 SZEGED, ARADI VÉRTANÚK TERE 1., HUNGARY

MTA-DE "LENDÜLET" FUNCTIONAL ANALYSIS RESEARCH GROUP, INSTITUTE OF MATHEMATICS, UNIVERSITY OF DEBRECEN, H-4010 DEBRECEN, P.O. BOX 12, HUNGARY

E-mail address: `gehergy@math.u-szeged.hu` or `gehergyuri@gmail.com`

URL: `http://www.math.u-szeged.hu/~gehergy/`